AD_____

Award Number: W81XWH-04-1-0081

TITLE: Molecular Aspects of Muscle Damage and Denervation with Public Access Tools

PRINCIPAL INVESTIGATOR: Eric P Hoffman, Ph.D.

CONTRACTING ORGANIZATION: Children's Research Institute
Washington DC 20010

REPORT DATE: December 2005

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01-12-2005	Final	1 Dec 2003 – 30 Nov 2005

4. TITLE AND SUBTITLE

Molecular Aspects of Muscle Damage and Denervation with Public Access Tools

5a CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-04-1-0081

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Eric P Hoffman, Ph.D.

Email: ehoffman@cnmcresearch.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Children's Research Institute
Washington DC 20010

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
Original contains colored plates: ALL DTIC reproductions will be in black and white.

14. ABSTRACT

Two Tasks were originally proposed; one was design, coding and implementation of PEPR (Public expression profiling resource). This was accomplished, and PEPR has become one of the most heavily used mRNA expression profiling (microarray) resources worldwide. We have populated with 7,000 microarray profiles, of which many are related to nerve and muscle damage and repair, and are of high relevance to health of the military. Of these, approximately 3,000 microarray profiles are in the public domain arm of PEPR, and have been co-submitted to NIH NCBI GEO through an automated data submission pipeline developed under the auspices of this award. Downloads of microarray data by the international research community is provided at no charge, and currently averages about 6,000 array downloads per month. This has effectively parallelized research on issues of importance the health of military recruits, including brain trauma, spinal cord injury, muscle exercise, muscle damage, and regenerative science. We also designed, coded and implemented HCE (Hierarchical Clustering Explorer). This powerful public domain software package has been downloaded by thousands of investigators, and facilitated thousands of research studies. Task 2 was focused on the proteomics of muscle atrophy and repair. Through support of the DoD award, we established proteomics expertise in the Research Center for Genetic Medicine. We also increased knowledge of the molecular pathways in muscle and nerve, damage. 21 publications in peer reviewed journals supported in whole or in part by the 2 yr DoD award. Some of these have been cited by the journals as "most highly accessed" of papers published.

15. SUBJECT TERMS
Muscle damage, muscle atrophy, remodeling, data warehouse, public access

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	20	19b. TELEPHONE NUMBER *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

INTRODUCTION:

The primary research thrust of the Department of Defense support was to design and implement a 2nd generation public access resource for microarray data. This was successfully accomplished with the launch of the Public Expression Profiling Resource (PEPR; http://pepr.cnmcresearch.org). PEPR includes an automated data submission pipeline to NCBI GEO, and also an API that automatically converts all projects to 5 probe set algorithms. The internal LIMS contains approximately 7,000 Affymetrix GeneChip profiles, with the majority (2,816) from human tissues. Of these, 2,827 are in the public domain through the PEPR public interface, as well as NCBI GEO. PEPR submissions account for 12% of all Affymetrix profiles in NCBI GEO, and our group is the #1 submitter of data to GEO.

Use of PEPR is quite high. PEPR averages 6,000 microarray downloads per month from the public, and many of the projects are highly relevant to the military (spinal cord damage, muscle damage and regeneration, brain trauma, exercise science).

We also developed the HCE public resource as free software for complex microarray data analyses through the DoD support. Over 4,600 scientists have downloaded the HCE software. This allows for visualization of complex multi-variant data sets.

17 publications in peer-reviewed journals were a direct outcome of the completed research, with many others published using the resources developed under the purview of the award.

BODY:

Two tasks were proposed in the original statement of work:

Task 1. Create a public access data warehouse for muscle with quality control and standard operating procedures, using a standardized platform, including muscle disease, exercise physiology, and plasticity following muscle data.

Task 2. Define the protein remodeling of the myofiber following two specific conditions known to damage muscle in humans; an atrophic stimulus (disuse following injury, denervation), and regeneration following injury. The injury model used involves "compensatory" changes that prevent further damage to the muscle; these compensatory remodeling events will be defined.

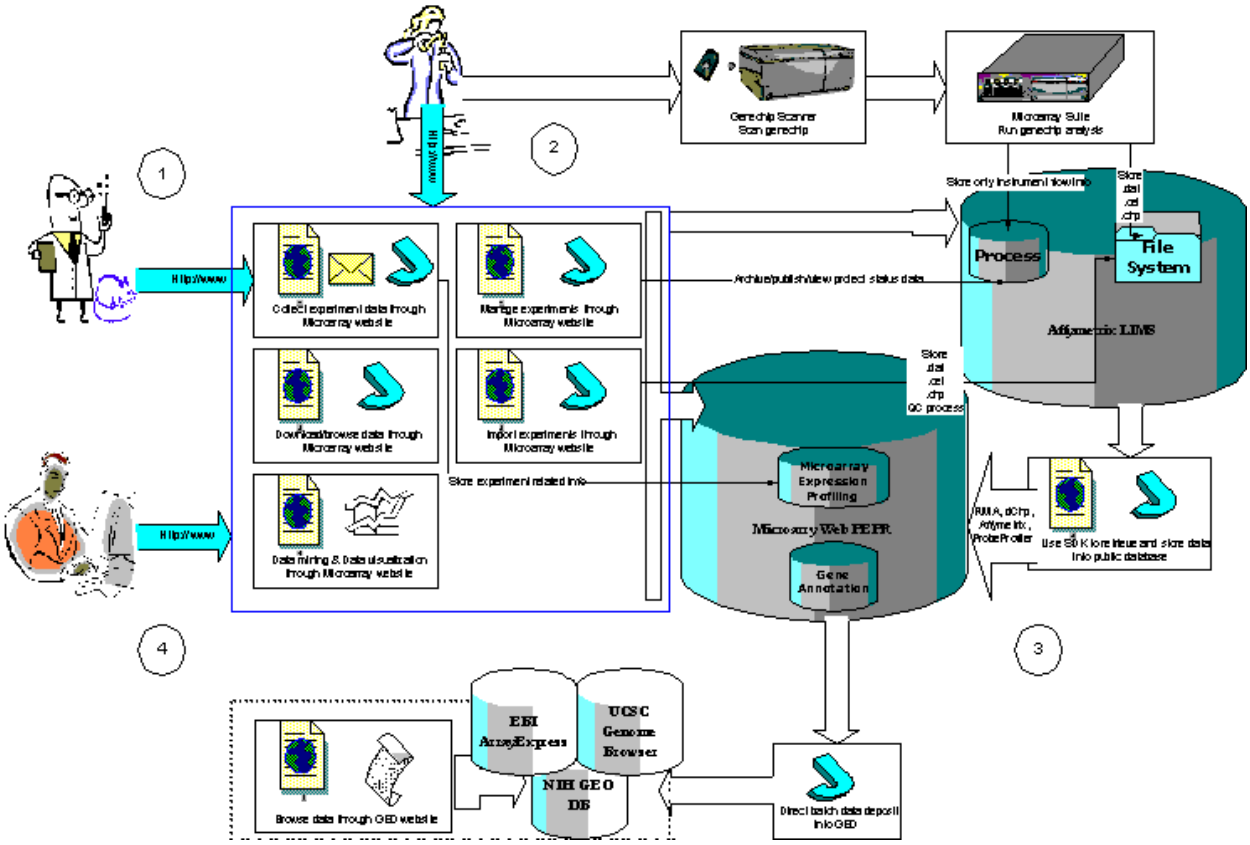
Overall, progress on Task 1 exceeded all plans and expectations. For example, the Statement of Work proposed 1,500 vertebrate microarrays to be done and implemented in the public resource, while we instead accomplished 7,000 profiles (see below). We also achieved an API for five distinct probe set algorithms, and achieved a public usage of approximately 6,000 downloads per month.

Progress on Task 2 required extensive development of proteomics resources and technologies. We had purchased a $850,000 ABI TOF/TOF unit for this Task from other funding (donations and Hospital contributions), but this unit was sent as a defective unit, and it took a full year before the unit became operational. Multiple attempts at resolving ubiquinated proteins following denervation were done, but these all failed despite extensive effort. Proteomics efforts then turned to both muscle metabolism (Hittel et al. 2005; Hittel et al. 2007), and ubiquination in statin myopathy (Urso et al. 2005). While these topics were not included in the original Statement of Work, they utilized the technologies proposed in the original grant.

A detailed description of progress on the two year award follows below:

Task 1. Create a public access data warehouse for muscle with quality control and standard operating procedures, using a standardized platform, including muscle disease, exercise physiology, and plasticity following muscle data.

We designed the PEPR resource according to the following schema:



The re-design described below enables rich meta-data search functions (i.e. search by experiment design type or animal model's age, sex); a web-interface data input system is used to capture experiment information. Unlike other currently utilized profiling packages, our web interface data input submission process offers great flexibility to obtain desired experiment meta-data (e.g. addition of experiment design type) for analysis and visualization. It provides a mechanism to enforce data input consistency and validation, and eliminates the current accessory tables and batch process to filter data. The data consistency expands the search and visualization capabilities.

Affymetrix GCOS operating system and AADM database is provided with all Affymetrix packages. However, rather than accessing the AADM database directly, our application utilizes the Affymetrix GCOS and GDAC SDK (software developer kit) to retrieve and parse experiment related data (e.g. .chp, .cel files). It preprocesses all the published chip files to improve the data download performance. It eliminates the existing process to transfer large sets of experiment data from lab database to public database. With GCOS and GDAC SDK, only a small subset of the data is extracted and placed in public database for analysis at any point in time. It also eliminates the AADM dependency (no need to change application if the AADM schema is changed). Indeed, the often-changing AADM schema resulted in chronic compatibility problems with the first generation PEPR resource.

PEPR also utilizes our newly implemented GEO submitted or update API's to submit new experiments or revised previously-published experiment data. PEPR incorporates a custom-designed Probe Profiler API (funded by a Department of Defense grant for PEPR to Dr. Hoffman), to offer four additional data algorithms (DCHP Diff, DCHP PMOnly, RMA, and PCA), in addition to the built-in MAS algorithm values for data analysis and visualization. Finally PEPR provides off-line batch data exportation that allows the researcher download/export a series of large data set while continuing to navigate the site. The generation of .chp, .dat and .cel data files is processed during off-peak hours.

Our previous design and implementation of PEPR was supported by an NHLBI Programs in Genomic Applications grant, and

an NINDS Spinal Cord Trauma grant (the latter the single NIH-award for this contract). While we have only very recently reported our initial implementation of PEPR (Almon et al. 2003; Chen et al. 2004), we feel our new re-design (funded by the Department of Defense and a R21 NHGRI grant) makes substantial improvements over our previous version, and any other dynamic query resource for massively parallel and multi-dimensional biological datasets available elsewhere.

The major improvements of PEPR while comparing the previous application include:

- proposal submission/approval workflow
- expanded search
- expanded data visualization
- data retrieval preprocess through GCOS and GDAC SDK
- GEO publishing addition and update
- Off-line batch data exportation

The major benefits of the PEPR while comparing the previous application:

- Workflow and central repository improves the collaboration between researchers and investigators.
- Enhanced search features offers better data sharing and navigation
- Enhanced visualization offers better assistances to researchers
- GCOS and GDAC SDK utilization eliminates the AADM dependency
- GEO publishing update completes the existing GEO publishing process (experiment addition and modification) through browser-based. It empowers the scientists to manage their own experiment data
- Off-line batch data exportation provides faster system response to researchers
- Data validation and consistency make database maintenance and operation easier
- OOD technology implementation make maintenance and future enhancement easier

The PEPR process architecture design and implementation

PEPR is a three-tier Java enterprise application, composed of a **Web Tier, Middle Tier and Back-End Tier** . A schematic of the overall design is provided on the next page of this text.

Web Tier

Web Tier includes a web server, a Tomcat application server and various web components which provide front end functionalities such as navigation, data browsing, data searching, project submission, project publishing, gene query tool and user notification. Most of web components interface transparently with PEPR back-end databases. This tier's interface allows users to trigger the middle tier application.

Middle Tier

The Middle Tier is integrated with several third party services, some of which we have purchased enterprise versions of pre-existing software, and others we wrote or contracted specifically for PEPR (Popchart, Lucene, Affymetrix SDK and Corimbia Probe Profiler SDK). It is designed to handle time-consuming processes such as Affymetrix data extraction, offline data downloading while allowing user to navigate the site without waiting the completion of the process. The Middle Tier applications require intense computing resources and are responsible for chart visualization generation, offline data download, metadata indexing for keyword search, NCBI GEO data submission; Affymetrix data file extraction and transformation, and Probe Profiler mixture of algorithm data generation.

Most of processes in this tier do not require synchronous response from the PEPR front-end. In addition to the conventional web click-and-wait applications features, PEPR allows user to submit the request without waiting the completion of the process while the process is guaranteed to be completed. To achieve this asynchronous operation in a reliable manner, an Open JMS queue server is introduced in PEPR implementation, and this serves to enhance the PEPR application functionalities. JMS is designed to handle the messages delivery between web components. When a user submits a request to download a large set of data in PEPR, a web component in Tomcat application server packages the user's request to a message and drops the message into the JMS Queue. The JMS Queue is responsible for receiving and delivering the message as a specialized router that looks at the message's address and delivers it to the appropriate parties (i.e. Offline Data Download process in the chart). The Offline Data Download process then parses and handles the download request. It continues to search and compress the requested data, and then send out the download URL notification to the user. During this process, the user does not have to wait for the lengthy file compression process completion. . The JMS Queue makes the batch download possible.

The importance of PEPR JMS Queue service:

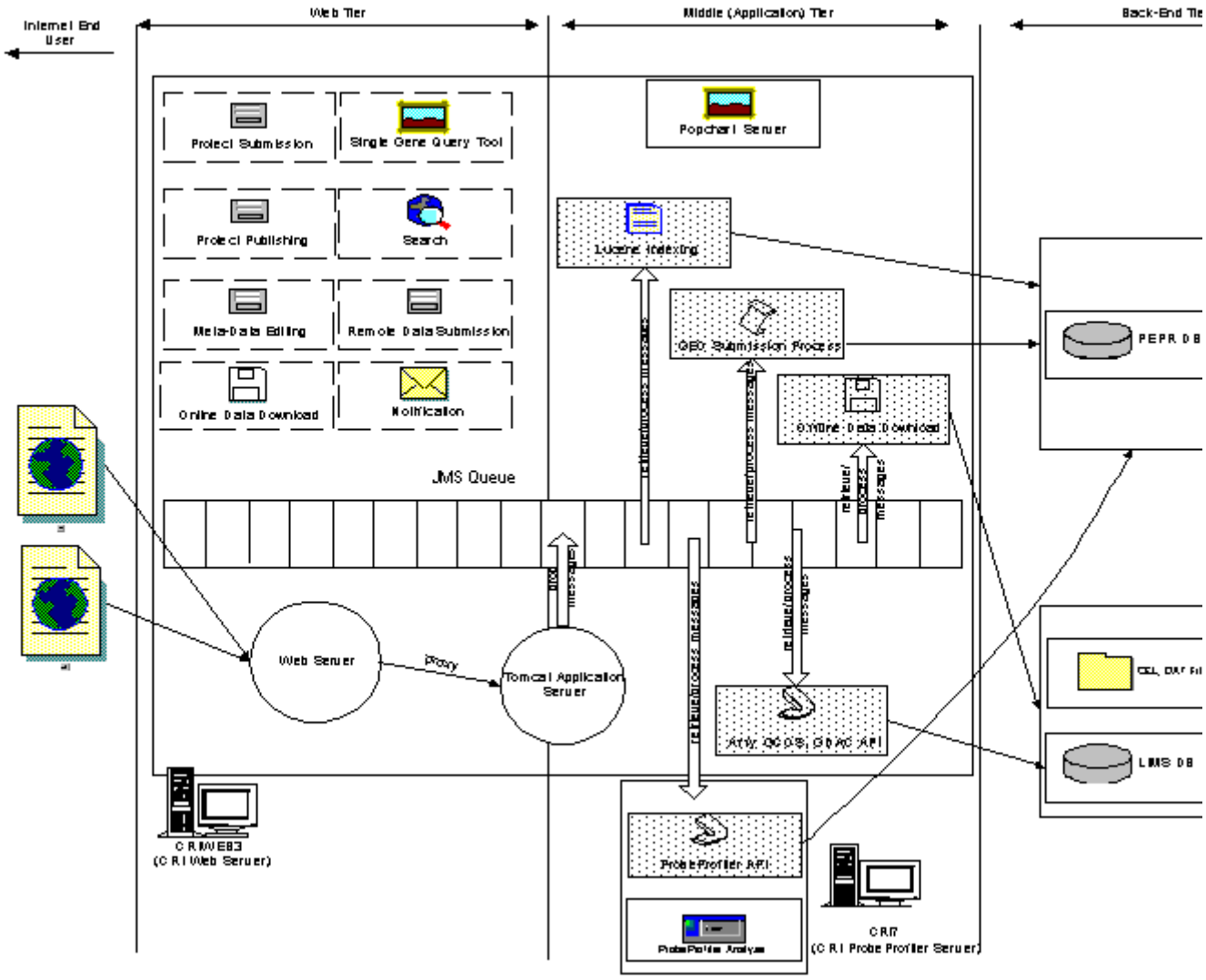
- Asynchronous communication: JMS Queue serves as an asynchronous communication channel between Web Tier and the Middle Tier components. When a PEPR administrator issues a GDAC data export command, the interface drops the message into JMS Queue and triggers the Affymetrix GDAC process, the process further loads data into the PEPR database while the administrator continues to perform other tasks.

- Reliable messaging communication: JMS Queue stores all the messages in Oracle database permanently. In the event of shutting down Middle Tier processes due to unexpected software failure, the JMS Queue continues to store and buffer the messages delivered from Tomcat application server. The JMS Queue then delivers the stored messages to the appropriate process when the Middle Tier applications restart. The persistence of JMS Queue provides PEPR high availability.

- Distributed computing: Probe Profiler API process requires intense computing resources. PEPR uses JMS Queue to distribute the computing resources to different server. JMS Queue is used to communicate with Probe Profiler API process (residing on CRI7) remotely. It allows the remote process to receive the messages and start its own calculation.

Sequence process control: Probe Profiler API is designed as single thread model; it can only process one request at a time. If more than one Probe Profiler processes are triggered at the same time, the second request would be dropped. JMS Queue can guarantee the arrival of the message and delivery of the message sequentially to the Probe Profiler API process on a first-come first served basis.

Figure. PEPR architecture.



Back-End Tier

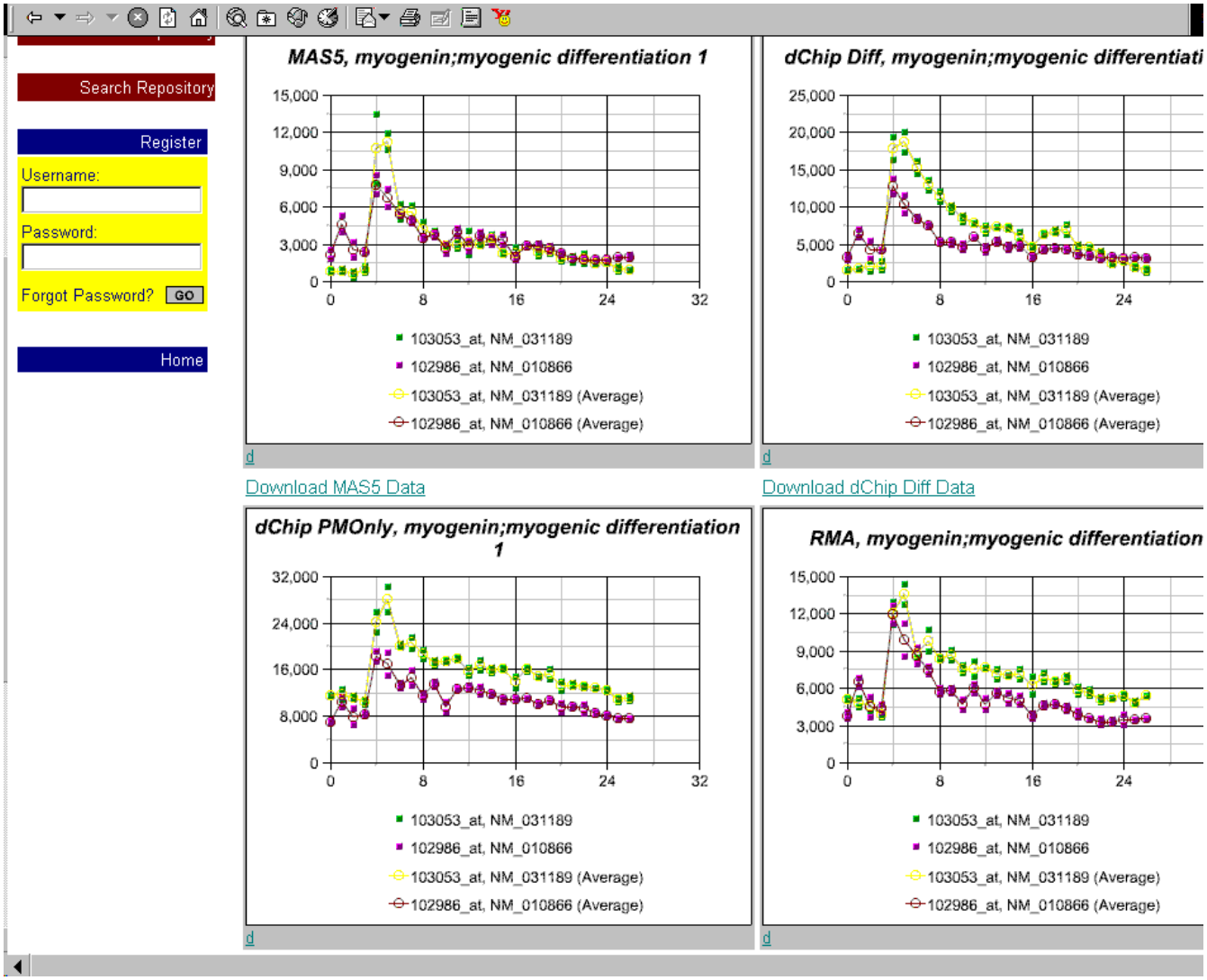
The Back-End Tier is composed of two databases; the PEPR DB and the Affymetrix LIMS DB. PEPR DB stores all sorts of metadata of projects and experiments alone with associated analysis value for real-time data mining purposes. The Affymetrix LIMS DB stores all Affymetrix expression profiling physical data and chipping process information.

We do not have adequate space to describe all the interfaces of PEPR, however we provide one screen snapshot of one interface (see following page). In this instance, we show dynamic query of a 27-time point time series project (see Zhao et al. 2002, 2003, 2004; Almon et al. 2003; Chen et al. 2004) (note that only 16 time points are shown in this example).

As can be seen, there are different specialized interfaces for the different types of users (left menu bars). Here, a web-based user has used a genome-browser type function to identify genes in the genome matching his/her query (e.g. “myogenic”), then used drop-down menus to select the specific gene and probe set that they wished to visualize. Multiple genes can be sent to be co-graphed; here, two myogenic factor genes were selected. The user can then define the probe set algorithm that should be visualized; here the user selected four of the available probe set algorithms. This dynamic query tool then extracts all data from the profiles, visualizes replicates, derives the averages of the replicates on the fly, graphs the genes relative to each other, provides mouse-overs showing all data behind that data point (including fold-change relative to time 0), and spreadsheets can be downloaded containing all data in the selected graphs. As can also be seen, different probe set

algorithms provide quite different interpretations, as we have previously reported (Seo et al. 2003).

Figure. Dynamic query web interface.



The total number of microarrays currently populating our internal LIMS is 7,000 with most from human specimens (Homo sapiens), as follows:

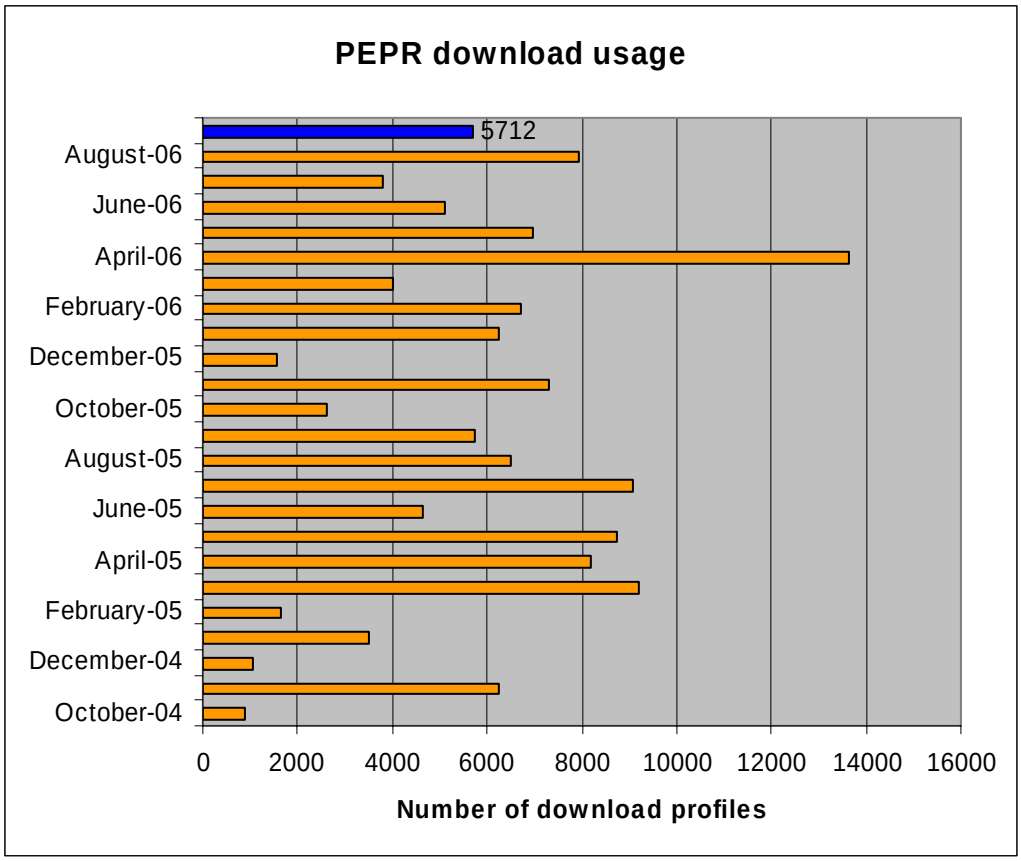
324 Other
2816 Homo sapiens
1906 Mus musculus
1823 Rattus norvegicus
131 Drosophila

Of these in the internal LIMS, 2,827 have been made public via PEPR, as follows:

Total of 2827
19 Canis domestica
810 Homo sapiens

809 Mus musculus
1189 Rattus norvegicus

Utilization of the public data source by external users is shown on the following graph.



Breaking down public usage by the top projects shows the following distributions. Note that the spinal cord trauma, and muscle regeneration are particularly relevant to the military.

# of downloads	# of profiles	Project Title	Tissue Type
30159	175	PGA Human CD4+ Lymphocytes	BLOOD
16236	239	Spinal Cord Trauma T9	SPINAL CORD
14351	242	Comparative profiling in 13 muscle disease groups	MUSCLE
13322	66	Muscle Regeneration	MUSCLE
6101	149	Spinal Cord Trauma Above T9	SPINAL CORD
5711	91	PGA Rat Liver Methylprednisolone	LIVER
4409	244	Spinal Cord Trauma Supraspinal Tracts	SPINAL CORD
3924	151	Spinal Cord Trauma Below T9	SPINAL CORD
3185	20	Human Glioblastoma	CANCER
2762	32	WKraus STRRIDE Study	MUSCLE
2613	57	Duchenne	MUSCLE
2436	40	PGA Human Cystic Fibrosis	BLOOD
2199	24	PGA Human Muscle Obese	MUSCLE

2105	23	Human Medulloblastoma	CANCER
1824	32	VSartorelli SMC differentiation	CELLS
1706	61	JNatale Murine Rat Brain Injury	BRAIN
1661	51	PGA Rat Muscle Methylprednisolone	MUSCLE
1285	63	PGA Rat Kidney Methylprednisolone	KIDNEY
1061	20	PGA Murine Glucose Metabolism	BRAIN
874	48	PGA Murine Lung Hyperoxia	LUNG
849	48	KEsser Rat Exercise	MUSCLE
810	24	PGA Murine IL-13 Asthma	LUNG
780	40	PGA Murine Airway Hyperresponsiveness	LUNG
776	52	PGA Murine Calories Restriction	LUNG
689	36	WSilk Macular Degeneration	EYE
657	80	DMD temporal profiling	MUSCLE
563	15	PGA Human Airway Hyperresponsiveness	LUNG
534	28	PGA Murine Lung Estrogen	LUNG
512	22	PGA Murine Pulmonary Fibrosis	LUNG
501	47	PGA Murine Lung Ragweed	LUNG
496	16	Spastic mouse	SPINAL CORD
460	30	Murine Neurofibromatosis	BRAIN
447	30	NINDS Rat Hippocampus Seizures	BRAIN
417	12	Skeletal Genome Anatomy Proj	BONE
368	24	PGA Murine Fibrillin-1 Deficient	BRAIN
336	10	FBooth MDX	MUSCLE
321	96	Spinal Cord Injury Murine Model	SPINAL CORD
308	44	Response of multiple genes to a chronic dose of corticosteroids in rat muscles	MUSCLE
296	30	PGA Murine Lung Hypertension	LUNG
289	15	Acute Quadriplegic Myopathy	MUSCLE
254	28	PGA Rat Necrotizing Enterocolitis	GUT
248	15	NINDS Rat Neuron Parkinsons	BRAIN
248	15	Pachman Juvenile Dermatomyositis	MUSCLE
219	11	NINDS Rat Epilepsy Diet	BRAIN
161	12	PRussell Human Glaucoma	MUSCLE
129	7	PGA Human Obstructive Pulmonary	MUSCLE
115	10	48h Immobilization in human	MUSCLE
108	12	PGA Human Broncial Epithelial	LUNG
108	8	PGA Murine Air Hyperpermability	LUNG
94	8	PGA Murine Lung Septation	LUNG
93	11	Hereditary Spastic Paraparesis	MUSCLE
84	12	PGA Rat Lung Seoul	LUNG
81	9	PGA Rat Lung Ventilation	LUNG
80	5	PGA Murine Cardiac Hypertrophy	HEART
78	24	p68 SMC differentiation	CELLS
66	6	Gastric Bypass Human Obese Muscle	MUSCLE
60	12	KNagaragu Murine Spleen	SPLEEN
50	23	Murine EDMD	MUSCLE
40	5	PGA Murine Goblet Cells	LUNG
36	5	PGA Dog Congestive Heart Failure	HEART
32	4	Normal Rat Muscle	MUSCLE
26	5	PGA Murine Alternatively Activated Macrophages	LUNG

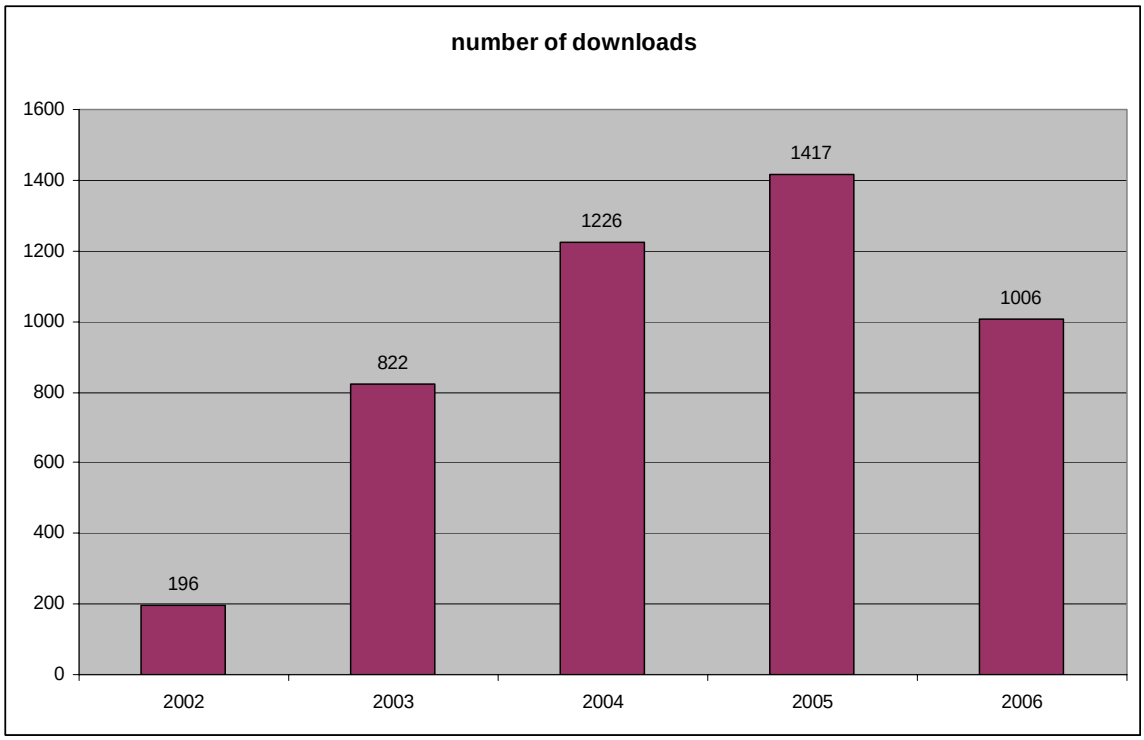
(AMM)
9 Myositis MUSCLE

130743 2783

Also done in the context of the DoD grant was design, implementation, and revisions of the HCE software. We modified the software to include power calculations for microarrays (Seo et al. 2006), and also investigated the effects of p value weighting in project-specific algorithm selection (Seo et al. 2004). Both of these papers were published in the top bioinformatics journal (Bioinformatics, Oxford University Press). We also published a solicited review article (Seo and Hoffman 2006); this become one of the most highly accessed papers in this popular journal in 2006 (thousands of downloads).

The HCE software is one of the most popular public domain packages for analysis of complex microarray data. Evidence of this is the number of downloads of the software from our web sites, as shown in the following graph.

HCE update for DoD grant



The PhD dissertation of Jinwook Seo was supported by the DoD grant, and it includes three compelling case studies of users in biology, statistics, and meteorology, who have produced published results. Their strong statements about how HCE changed their work make for engaging reading, e.g. "extremely useful … Typically gaining this type of information using statistics packages is very time consuming." Another strong support for this contribution is that the research related to HCE led to several journal publications.

Five papers were published in major biology and bioinformatics journals. Three papers were published in top journals on information visualization and human computer interaction. HCE works were also presented in three international conferences. In addition, HCE has been used by many researchers around world and cited by them in many journal papers mostly in biology but also in other disciplines.

Another major contribution of this work is to promote evaluation methods appropriate for information visualization and other creativity support tools. Since controlled experimental studies are not likely to capture the experience of domain specialists working on deep problems, Jinwook Seo conducted indepth participant observations and interviews of researchers in molecular biology, statistics, and meteorology over 6-week periods. In addition, he collected email survey data from 57 serious users to assess which features were most helpful.

Figure. The cities from which the most visitors come to HCE homepage in a month of July 2006

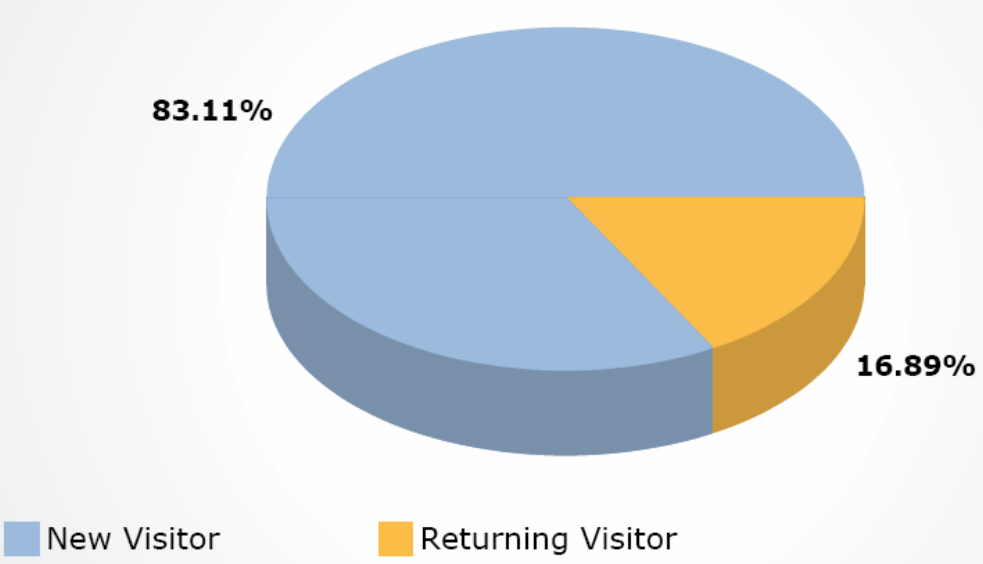


Figure The number of first-time visits and returning visits in a month of July 2006

Overall, the DoD supported developments of both PEPR and HCE leveraged a large number of publications in both bioinformatics and muscle biology, all relevant to the original statement of work. There have been publications on pharmacogenomics (steroids) (Almon et al. 2003), muscle disease (Bakay et al. 2006; Chen et al. 2005; Melcon et al. 2006; Molon et al. 2004; Urso et al. 2005), exercise physiology (Chen et al. 2003; Hittel et al. 2007), muscle molecular biology (Caretti et al. 2006; Iezzi et al. 2004;), dimorphism (Lamason et al. 2006), muscle regeneration (Zhao et al. 2006a; Zhao et al. 2006b; Zhao and Hoffman 2004; Zhao et al. 2003) and diabetes (Hittel et al. 2005). Many additional publications have been on bioinformatics developments, all directly supported by the DoD grant (Chen et al. 2004; Seo et al. 2004; Seo et al. 2006a; Seo et al. 2006b).

For the second major Statement of Work, namely proteomics development and applications, this proved quite problematic at a number of levels. First, we took delivery on an $850,000 state-of-the-art proteomics unit, the ABI TOF/TOF to carry out this aim. This was supported by donations and hospital contributions as matching funds for the DoD grant. This was a “pre-release” unit, and unfortunately, the unit was shipped with incorrect lasers and collision chambers, and did not function. More unfortunately, the company refused to acknowledge that the unit was defective, and took a full year before replacing adequate parts to make the unit functional. Even more unfortunately, ABI refused to provide any compensation for the lost year of work on the defective unit.

We also ran into trouble with the isolation of ubiquinated proteins; a key first step in the proteomic characterizations. Ubiquinated forms of proteins have a very short half life, making them highly transient and unstable by products. This proved beyond our ability to isolate effectively, despite extensive attempts by an experienced post-doc in proteomics covered by the DoD grant (Dr. Kristy Brown). In an attempt to achieve related work on this aim, we arranged a collaboration with Fred Goldberg of Harvard Medical School, including a visit to his laboratory. He agreed to send a series of RNA samples related to atrophic conditions (the focus of this section of the statement of work for the DoD grant), that we would then expression profile, and apply the advanced bioinformatics tools developed above. Unfortunately, this shipment was lost for a number of days by FedEx, and arrived thawed, and all RNA degraded.

While the progress on the second part of the Statement of Work (proteomics of muscle atrophy) was disappointing, we did end up laying the groundwork for a thriving proteomics group that has had success on a series of other projects. These have included 8 publications on a variety of topics by collaborator Yetrib Hathout, and some focused papers on muscle proteomics (Hittel et al. 2005; Hittel et al. 2007). One particularly interesting paper is focused on the molecular definition of the neuromuscular junction; a key cellular subspecialization where the motor neuron hits the nerve, and a frequent target of biological warfare, as well as muscle atrophy/damage (Nazarian et al. 2007).

KEY RESEARCH ACCOMPLISHMENTS: Bulleted list of key research accomplishments emanating from this research.

- Design, coding and implementation of PEPR (Public expression profiling resource), one of the most heavily used mRNA expression profiling (microarray) resources worldwide.
- Population of PEPR with over 3,000 microarray profiles, many on projects of high relevance to the military (muscle exercise, damage, brain damage, nerve damage and repair).
- Downloads of over 60,000 profiles by researchers worldwide, effectively parallelizing research on issues of importance the health of military recruits.
- Design, coding and implementation of HCE (Hierarchical Clustering Explorer). This has been downloaded by thousands of investigators, and facilitated thousands of research studies.
- Establishment of proteomics expertise in the Research Center for Genetic Medicine
- Increased knowledge of the molecular pathways in muscle and nerve, damage.
- 21 publications in peer reviewed journals supported in whole or in part by the DoD award. Some of these have been cited by the journals as “most highly accessed” of papers published.

REPORTABLE OUTCOMES: Provide a list of reportable outcomes that have resulted from this research to include:

Manuscripts: 21 publications. See References for complete list (all references are those supported by this award).

Presentations: Dozens of invited presentations on muscle disease, damage, repair, proteomics, expression profiling, and bioinformatics. A partial list follows:

- Neuroscience Seminar, University of California San Francisco, San Francisco, CA
- American Association of Allergy, Auto-immunity, and Immunology (AAAAI), San Francisco, CA
- Affymetrix Core Directors’ Meeting, Speaker, New Orleans, LA
- Departmental Seminar, Department of Computer Science, “Signal/noise assessment in microarray data”, University of Maryland, College Park, MD
- Bio-informatics Symposium, Buffalo Center for Biomedical Computing, “Expression profiling to define biochemical pathways”, SUNY Buffalo, NY

- Nobel Symposium on Inflammatory Myopathies, “Biochemical pathways in muscle disease”, Karolinska Institute, Stockholm, Sweden
- Molecular basis of muscle atrophy symposium, University of Massachusetts Amherst, MA.
- University of Pennsylvania, Muscle Research Institute, Philadelphia PA.
- Merck Research Laboratories, Research Seminar Series, Muscle pathology group, New Jersey.
- Case Western Reserve University, MD/PhD retreat, Keynote speaker, SNP associations in body type, Columbus, OH
- Virginia Neurological Association, Molecular Diagnosis of Muscular Dystrophy, Hot Springs, West Virginia
- George Washington University, Keck proteomics symposium, Washington DC.
- University of Pennsylvania, Mental Retardation and Developmental Disabilities Seminar Series (MRDDRC), Philadelphia PA
- Affymetrix Core Directors’ Meeting, Featured Speaker, New Orleans, LA
- Novartis Research Seminar, The genetics of type II diabetes, Boston, MA
- North Carolina National Society for Genetic Counselors, Keynote speaker, Wake Forest University, Winston-Salem, NC
- Muscle regeneration and stem cells, Transcriptional pathways in muscle regeneration, FASEB meeting, Tucson AZ
- Nuclear architecture and human disease, Molecular basis of Emery-Dreifuss Muscular Dystrophy, ACSB, Des Moines, IA
- US Anti-Doping Agency, Genetics of muscle performance, Chicago, IL

Patents and Licenses; Cell lines; Tissue repositories: N/A

Informatics:

PEPR (http://pepr.cnmcresearch.org)

HCE (http://www.dcchildrens.com/cnmcresearch/bioinformatics/power/power.html)

SAS server (http://sas.cnmcresearch.org)

Funding applied for (and received) based on the work supported by this award:

3R01 NS29525-13 (Hoffman) 01/01/91-11/30/10
Improved Diagnosis of the Muscular Dystrophies
The theme of this grant is to determine the molecular basis of the muscular dystrophies, using both candidate gene/protein approaches, and genome-wide discovery approaches.

W81XWH-05-0334 (Hoffman) 10/1/04-9/31/07
Molecular mechanisms of corticosteroid action on muscle physiology.
The goal of this grant is to determine the molecular basis of the enigmatic beneficial effects of chronic corticosteroid on Duchenne muscular dystrophy muscle. The hypothesis to be tested is that corticosteroids have influence three major pathways: anti-inflammatory, catabolic (via AKT1 signaling), and anabolic (via transcriptional responses and metabolic integration).

W81XWH-05-1-0616 (Hoffman) 9/15/05-9/14/08
Muscle Research Consortium (Program Project): Duchenne muscular dystrophy.
This program project involves four research projects: Development of high throughput drug screening assays (Miceli, UCLA), mechanisms of muscle atrophy (Sweeney, U of Penn), oligonucleotide directed splicing approaches (Lu, Carolinas Med Inst), and muscle stem cell biology (Partridge, Children’s DC). Two cores are funded; Administrative (Hoffman, CNMC), and Mouse Functional Testing Core (Nagaraju, CNMC).

5R24HD050846-02 (Hoffman) 10/1/05-9/30/10
Integrated molecular core for rehabilitation medicine.
This is a core facility to provide DNA, mRNA, proteomics, and database services to grantees of the NICHD Medical Rehabilitation Research Center.

1U54HD053177-01A1 (Hoffman) 10/1/05-9/30/10
Wellstone Muscular Dystrophy Center: Children's National Medical Center
This Center grant includes three projects and three cores. Project 1 (Hoffman PI, Escolar Co-PI) is a SNP association study in Duchenne muscular dystrophy patients, looking at both corticosteroid responsiveness and natural history. Project 2 (Chen PI, Nagaraju Co-PI) looks at NFkB and TGFbeta cascades during the progression of Duchenne dystrophy in human and mouse models. Project 3 (Partridge PI) looks at stem cell populations in muscle, and the effect of IGF-1 on determination and proliferation. Core A (Hoffman PI) is Administrative, Core B (Human Clinical Core; Escolar PI) supports the CINRG clinical trial network, and Core C (Bioinformatics Core; Chen PI) provides computing, statistical, and bioinformatics support.

RO1 NS40606-05 (Hoffman) NIH NINDS/NIAMS/NIA 6/1/01-7/31/11
Functional SNPs Associated with Muscle Size and Strength
The specific aims of this competitive renewal are to continue analysis of this pre-existing cohort of subjects, both with regards to phenotyping (completion of volumetric studies, extension to individual muscle groups and focal size changes, development of a public access resource to the data) (Aim 1), and genotyping (new loci, validation of existing associations through testing of GUSTO and Health ABC cohorts, and extension of haplotypes) (Aim 2). Aim 3 is focused on systematically defining the effects of the robust AKT1 associations with functional consequences of the component SNPs on AKT1 gene promoter function, and a potential Zinc finger transcript unit upstream of AKT1.

<u>CORE FUNCTIONS:</u>

NIH NCRR
General Clinical Research Center (Tuchman) 12/1/00-11/30/09
Genetics Core Laboratory (Hoffman)
The goal of the Genetics Core Laboratory is to provide genotyping and expression profiling services to the PCRC.

1P30HD40677-01 (Tuchman) 8/1/01-7/31/11
NIH, NICHD
MRDDRC at Children's National Medical Center
Molecular Genetics Core (Hoffman)
The main goal of this project is the operation of a center of excellence for research and training in the area of mental retardation and developmental disabilities in Washington, D.C. The goal of the Molecular Genetics Core is to provide nucleic acids research support to all members of the Center, including expression profiling, sequencing, and genotyping.

CONCLUSION:

In conclusion, we have made outstanding progress in Task 1. Namely we have created a public access data warehouse for muscle with quality control and standard operating procedures, using a standardized platform, including muscle disease, exercise physiology, and plasticity following muscle data. Our performance met all statement of work objectives, and the popularity of the software and database tools by the international research community vastly exceeded our expectations. One method of evaluating this is the “leverage” provided by the award. Given 60,000 downloads of microarray data from PEPR alone, the cost of generating this data by each individual scientist using it would have exceeded $60 million. Thus, there was greater than a 60-fold leverage, where a $1 million investment by DoD leveraged an additional $60 million of scientific activity and research. We know that our data on muscle and nerve damage and plasticity is among the most highly utilized in NIH NCBI GEO as well, so the $60 million is likely a conservative estimate.

Our progress on Task 2 was more measured. Task 2 was to define the protein remodeling of the myofiber following two specific conditions known to damage muscle in humans; an atrophic stimulus (disuse following injury, denervation), and regeneration following injury. The injury model used involves “compensatory” changes that prevent further damage to the muscle; these compensatory remodeling events will be defined. The proteomic part of this Task was beset with technical difficulties, both due to very expensive machinery that took over a year to become functional, and difficulties in isolating very unstable ubiquitinated proteins.

Our publication record is considered outstanding, with 21 publications in peer reviewed journals related to the original statement of work. We also presented dozens of invited lectureships, and cited work done under the auspices of the DoD grant at each of these presentations.

REFERENCES:

The following manuscripts were published, and were supported wholly or in part by the DoD award.

Almon RR, Chen J, Snyder G, DuBois DC, Jusko WJ, Hoffman EP. In vivo multi-tissue corticosteroid microarray time series available online at Public Expression Profile Resource (PEPR). **Pharmacogenomics.** 2003 Nov;4(6):791-9.

Bakay M, Wang Z, Melcon G, Schiltz L, Xuan J, Zhao P, Sartorelli V, Seo J, Pegoraro E, Angelini C, Shneiderman B, Escolar D, Chen YW, Winokur ST, Pachman LM, Fan C, Mandler R, Nevo Y, Gordon E, Zhu Y, Dong Y, Wang Y, Hoffman EP. Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb-MyoD pathways in muscle regeneration. **Brain.** 2006 Apr;129(Pt 4):996-1013.

Caretti G, Schiltz RL, Dilworth FJ, Di Padova M, Zhao P, Ogryzko V, Fuller-Pace FV, Hoffman EP, Tapscott SJ, Sartorelli V. The RNA helicases p68/p72 and the noncoding RNA SRA are coregulators of MyoD and skeletal muscle differentiation. **Dev Cell.** 2006 Oct;11(4):547-60.

Chen J, Zhao P, Massaro D, Clerch LB, Almon RR, DuBois DC, Jusko WJ, Hoffman EP. The PEPR GeneChip data warehouse, and implementation of a dynamic time series query tool (SGQT) with graphical interface. **Nucleic Acids Res**. 2004 Jan 1;32(Database issue):D578-81.

Chen YW, Hubal MJ, Hoffman EP, Thompson PD, Clarkson PM. Molecular responses of human muscle to eccentric exercise. **J Appl Physiol.** 2003 Dec;95(6):2485-94.

Chen YW, Nagaraju K, Bakay M, McIntyre O, Rawat R, Shi R, Hoffman EP. Early onset of inflammation and later involvement of TGFbeta in Duchenne muscular dystrophy. **Neurology**. 2005 Sep 27;65(6):826-34.

Hittel DS, Hathout Y, Hoffman EP, Houmard JA. Proteome analysis of skeletal muscle from obese and morbidly obese women. **Diabetes**. 2005 May;54(5):1283-8.

Hittel DS, Hathout Y, Hoffman EP. Proteomics and systems biology in exercise and sport sciences research. **Exerc Sport Sci Rev.** 2007 Jan;35(1):5-11.

Iezzi S, Di Padova M, Serra C, Caretti G, Simone C, Maklan E, Minetti G, Zhao P, Hoffman EP, Puri PL, Sartorelli V. Deacetylase inhibitors increase muscle cell size by promoting myoblast recruitment and fusion through induction of follistatin. **Dev Cell**. 2004 May;6(5):673-84.

Lamason R, Zhao P, Rawat R, Davis A, Hall JC, Chae JJ, Agarwal R, Cohen P, Rosen A, Hoffman EP, Nagaraju K. Sexual dimorphism in immune response genes as a function of puberty. **BMC Immunol.** 2006 Feb 22;7:2.

Melcon G, Kozlov S, Cutler DA, Sullivan T, Hernandez L, Zhao P, Mitchell S, Nader G, Bakay M, Rottman JN, Hoffman EP, Stewart CL. Loss of emerin at the nuclear envelope disrupts the Rb1/E2F and MyoD pathways during muscle regeneration. **Hum Mol Genet.** 2006 Feb 15;15(4):637-51.

Molon A, Di Giovanni S, Chen YW, Clarkson PM, Angelini C, Pegoraro E, Hoffman EP. Large-scale disruption of microtubule pathways in morphologically normal human spastin muscle. **Neurology**. 2004 Apr 13;62(7):1097-104.

Nazarian J, Hathout Y, Hoffman EP. The Proteome Survey of an Electricity-Generating Organ (Torpedo californica electric organ). **Proteomics**, *in press.*

Seo J, Bakay M, Chen YW, Hilmer S, Shneiderman B, Hoffman EP. Interactively optimizing signal-to-noise ratios in expression profiling: project-specific algorithm selection and detection p-value weighting in Affymetrix microarrays. **Bioinformatics**. 2004 Nov 1;20(16):2534-44.

Seo J, Gordish-Dressman H, Hoffman EP. An interactive power analysis tool for microarray hypothesis testing and generation. **Bioinformatics**. 2006a Apr 1;22(7):808-14.

Seo J, Hoffman EP. Probe set algorithms: is there a rational best bet? **BMC Bioinformatics.** 2006b Aug 30;7:395.

Urso ML, Clarkson PM, Hittel D, Hoffman EP, Thompson PD. Changes in ubiquitin proteasome pathway gene expression in skeletal muscle with exercise and statins. **Arterioscler Thromb Vasc Biol.** 2005 Dec;25(12):2560-6.

Zhao P, Caretti G, Mitchell S, McKeehan WL, Boskey AL, Pachman LM, Sartorelli V, Hoffman EP. Fgfr4 is required for effective muscle regeneration in vivo. Delineation of aMyoD-Tead2-Fgfr4 transcriptional pathway. **J Biol Chem.** 2006a Jan 6;281(1):429-38.

Zhao P, Hoffman EP. Musculin isoforms and repression of MyoD in muscle regeneration. **Biochem Biophys Res Commun.** 2006b Apr 14;342(3):835-42.

Zhao P, Hoffman EP. Embryonic myogenesis pathways in muscle regeneration. **Dev Dyn**. 2004 Feb;229(2):380-92.

Zhao P, Seo J, Wang Z, Wang Y, Shneiderman B, Hoffman EP. In vivo filtering of in vitro expression data reveals MyoD targets. **C R Biol.** 2003 Oct-Nov;326(10-11):1049-65.

APPENDICES: Attach all appendices that contain information that supplements, clarifies or supports the text. Examples include original copies of journal articles, reprints of manuscripts and abstracts, a curriculum vitae, patent applications, study questionnaires, and surveys, etc.

N/A

SUPPORTING DATA: All figures and/or tables shall include legends and be clearly marked with figure/table numbers.

N/A